

A Generalized Cheeger Inequality *

Ioannis Koutis
ioannis.koutis@upr.edu

Gary Miller
glmiller@cs.cmu.edu

Richard Peng
rpeng@mit.edu

December 19, 2014

Abstract

The generalized conductance $\phi(G, H)$ between two graphs G and H on the same vertex set V is defined as the ratio

$$\phi(G, H) = \min_{S \subseteq V} \frac{\text{cap}_G(S, \bar{S})}{\text{cap}_H(S, \bar{S})},$$

where $\text{cap}_G(S, \bar{S})$ is the total weight of the edges crossing from S to $\bar{S} = V - S$. We show that the minimum generalized eigenvalue $\lambda(L_G, L_H)$ of the pair of Laplacians L_G and L_H satisfies

$$\lambda(L_G, L_H) \geq \phi(G, H)\phi(G)/8,$$

where $\phi(G)$ is the usual conductance of G . A generalized cut that meets this bound can be obtained from the generalized eigenvector corresponding to $\lambda(L_G, L_H)$. The inequality complements a result of Trevisan [Tre13] which shows that $\phi(G)$ cannot be replaced by $\Theta(\phi(G, H))$ in the above inequality, unless the Unique Games Conjecture is false.

1 The Cheeger inequality

Let $G = (V, E, w)$ be a connected weighted graph. For $v \in V$ and $S \subseteq V$ we let

$$\text{vol}(v) = \sum_{(v,w) \in E} w(v, w) \quad \text{and} \quad \text{vol}(S) = \sum_{v \in S} \text{vol}(v).$$

We will call $\text{vol}(v)$ the degree of vertex v . We also denote by $\text{cap}(S, \bar{S})$ the total weight of edges with exactly one endpoint in S and one endpoint in \bar{S} . The **conductance** of G is defined as

$$\phi(G) = \min_{S \subseteq V} \frac{\text{cap}(S, \bar{S})}{\min\{\text{vol}(S), \text{vol}(\bar{S})\}}.$$

The Laplacian of G is defined by

$$L(u, v) = -w(u, v) \quad \text{and} \quad L(u, u) = \sum_{v \neq u} w(u, v).$$

*Partially supported by the National Science Foundation under grant numbers CCF-1018463 and CCF-1149048.

The normalized Laplacian \tilde{L} of G is the matrix $D^{-1/2}LD^{-1/2}$ where D is the diagonal of L . It is well understood that the normalized Laplacian is positive semi-definite with a unique zero eigenvalue. If λ_2 is its second eigenvalue λ_2 , then the Cheeger inequality relates it to $\phi(G)$ as follows:

$$\lambda_2 \geq \phi(G)^2/2.$$

At least one proof of the Cheeger inequality due to Mihail [Mih89] actually shows something stronger, namely that for any vector $y \perp \text{Null}(\tilde{L}_G)$, we can find a set S_y such that

$$y^T \tilde{L}_G y > \left(\frac{\text{cap}_G(S_y, \bar{S}_y)}{\min\{\text{vol}(S_y), \text{vol}(\bar{S}_y)\}} \right)^2 / 2. \quad (1.1)$$

The cut can be found by letting S_y to consist of the vertices corresponding to the k smallest entries of y , for some $1 \leq k \leq n$.

2 Generalized cuts for graph pairs

We define the **generalized conductance** between G and H as follows:

$$\phi(G, H) = \min_{S \subseteq V} \frac{\text{cap}_G(S, \bar{S})}{\text{cap}_H(S, \bar{S})}.$$

To motivate this definition, we observe that the sparsest cut problem can be captured within a factor of 2 as a generalized cut problem¹ between two graphs. To this end let us define the **demand graph** $D_G = (V, E', w')$ with every edge being present in E' and the weights specified by

$$w'(u, v) = \frac{\text{vol}(u)\text{vol}(v)}{\text{vol}(V)}.$$

Let $S \subseteq V$. Observe that by construction we have

$$\text{cap}_{D_G}(S, \bar{S}) = \frac{\text{vol}(S)\text{vol}(\bar{S})}{\text{vol}(V)}.$$

Note now that

$$\min\{\text{vol}(S), \text{vol}(\bar{S})\} \geq \frac{\text{vol}(S)\text{vol}(\bar{S})}{\text{vol}(V)} \geq \min\{\text{vol}(S), \text{vol}(\bar{S})\}/2.$$

From this it can be seen that

$$\frac{\phi(G)}{2} \leq \phi(G, D_G) \leq \phi(G).$$

A number of other problems can be viewed as generalized cut problems.

For example, consider the **isoperimetric number** defined by:

$$h(G) = \min_{S \subseteq V} \frac{\text{cap}_G(S, \bar{S})}{\min\{|S|, |\bar{S}|\}}.$$

¹sometimes called the non-uniform sparsest cut problem.

If K_n is the complete graph on n vertices with edges weighted by $1/n$, i.e. the identity over the space of sets orthogonal to the constant vectors, it can be verified that we have

$$\frac{\phi(G, K_n)}{2} \leq h(G) \leq \phi(G, K_n).$$

Another example is the min s - t cut problem which looks for a cut of minimum value among all possible cuts that separate s and t . If we denote that value by $\mu_{s,t}$, and let $G_{s,t}$ be the Laplacian of the edge (s, t) , we have

$$\mu_{s,t} = \phi(G, G_{s,t}).$$

2.1 Cuts and Laplacians

The value of a cut between S and \bar{S} can be expressed in terms of the graph Laplacian as:

$$\text{cap}_G(S, \bar{S}) = x_S^T L_G x_S,$$

where x_S is characteristic vector of S , i.e. the vector with ones in its entries corresponding to S and zeros in all other entries.

It follows that the generalized conductance can be expressed as an optimization problem over the discrete 0-1 vectors:

$$\phi(G, H) = \min_{x \in \{0,1\}^n} \frac{x^T L_G x}{x^T L_H x}.$$

One can relax this discrete problem over the real numbers:

$$\lambda(G, H) = \min_{\substack{x \in \mathbb{R}^n \\ x^T d = 0}} \frac{x^T L_G x}{x^T L_H x}.$$

Here d is the vector containing the degrees of the vertices in G . The constraint $x^T d = 0$ can be considered as a ‘normalization’ constraint that fixes one representative for all vectors of the form $x_c = y + c1$ that achieve the same ratio. That is, the constraint doesn’t change the minimum value of the ratio and we add it here because it will be useful in the sequel. It is well known that $\lambda(G, H)$ is the first non-trivial generalized eigenvalue for the problem $L_G x = \lambda L_H x$.

Note that the minimum eigenvalue of the normalized Laplacian can also be seen as the first non-trivial generalized eigenvalue for the problem $Lx = \lambda Dx$ and thus as a continuous relaxation of the corresponding optimization problem. We aim to prove a similar characterization for the generalized conductance of any pair of graphs.

3 Generalized Cheeger Inequality

We begin with two Lemmas.

Lemma 1. *For all $a_i, b_i > 0$ we have*

$$\frac{\sum_i a_i}{\sum_i b_i} \geq \min_i \left\{ \frac{a_i}{b_i} \right\}.$$

Lemma 2. Let G be a graph, d be the vector containing the degrees of the vertices, and D be corresponding diagonal matrix. For all vectors x where $x^T d = 0$ we have

$$x^T D x = x^T L_{D_G} x,$$

where D_G is the demand graph for G .

Proof. Let d be the vector consisting of the entries along the diagonal of D . By definition, we have

$$L_{D_G} = D - \frac{dd^T}{\text{vol}(V)}.$$

The lemma follows. □

We prove the following theorem.

Theorem 1. Let G and H be any two weighted graphs and D be the vector containing the degrees of the vertices in G . For any vector x such that $x^T d = 0$, we have

$$\frac{x^T L_G x}{x^T L_H x} \geq \phi(G, D_G) \cdot \phi(G, H)/4,$$

where D_G is the demand graph of G

Let V^- denote the set of u such that $x_u \leq 0$ and V^+ denote the set such that $x_u > 0$. Then we can divide E_G into two sets: E_G^{same} consisting of edges with both endpoints in V^- or V^+ , and E_G^{dif} consisting of edges with one endpoint in each. In other words:

$$\begin{aligned} E_G^{\text{dif}} &= \delta_G(V^-, V^+), \text{ and} \\ E_G^{\text{same}} &= E_G \setminus E_G^{\text{dif}}. \end{aligned}$$

We also define E_H^{dif} and E_H^{same} similarly.

We first show a lemma which is identical to one used in the proof of Cheeger's inequality [Chu97]:

Lemma 3. Let G and H be any two weighted graphs on the same vertex set V partitioned into V^- and V^+ . For any vector x we have

$$\frac{\sum_{uv \in E_G^{\text{same}}} w_G(u, v) |x_u^2 - x_v^2| + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2)}{x^T L_H x} \geq \frac{\phi(G, H)}{2}.$$

Proof. We begin with a few algebraic identities:

Note that $2x_u^2 + 2x_v^2 - (x_u - x_v)^2 = (x_u + x_v)^2 \geq 0$ gives:

$$(x_u - x_v)^2 \leq 2x_u^2 + 2x_v^2.$$

Also, suppose $uv \in E_H^{\text{same}}$ and without loss of generality that $|x_u| \geq |x_v|$. Then letting $y = |x_u| - |x_v|$, we get:

$$\begin{aligned} |x_u^2 - x_v^2| &= (|x_v| + y)^2 - |x_v|^2 \\ &= y^2 + y|x_v| \\ &\geq y^2 = (x_u - x_v)^2. \end{aligned}$$

The last equality follows because x_u and x_v have the same sign.

We then use the above inequalities to decompose the $x^T L_H x$ term.

$$\begin{aligned}
x^T L_H &= \sum_{uv \in E_H^{same}} w_H(u, v) (x_u - x_v)^2 + \sum_{uv \in E_H^{dif}} w_H(u, v) (x_u - x_v)^2 \\
&\leq \sum_{uv \in E_H^{same}} w_H(u, v) (x_u - x_v)^2 + \sum_{uv \in E_H^{dif}} w_H(u, v) (2x_u^2 + 2x_v^2) \\
&\leq 2 \left(\sum_{uv \in E_H^{same}} w_H(u, v) (x_u - x_v)^2 + \sum_{uv \in E_H^{dif}} w_H(u, v) (x_u^2 + x_v^2) \right) \\
&\leq 2 \left(\sum_{uv \in E_H^{same}} w_H(u, v) |x_u^2 - x_v^2| + \sum_{uv \in E_H^{dif}} w_H(u, v) (x_u^2 + x_v^2) \right). \tag{3.1}
\end{aligned}$$

We can now decompose the summation further into parts for V^- and V^+ :

$$\begin{aligned}
&\sum_{uv \in E_G^{same}} w_G(u, v) |x_u^2 - x_v^2| + \sum_{uv \in E_G^{dif}} w_G(u, v) (x_u^2 + x_v^2) \\
&= \sum_{u \in V^-, v \in V^-} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_u^2 \\
&\quad + \sum_{u \in V^+, v \in V^+} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_v^2.
\end{aligned}$$

Doing the same for $\sum_{uv \in E_H^{same}} w_H(u, v) |x_u^2 - x_v^2| + \sum_{uv \in E_H^{dif}} w_H(u, v) (x_u^2 + x_v^2)$ we get:

$$\begin{aligned}
&\frac{\sum_{uv \in E_G^{same}} w_G(u, v) |x_u^2 - x_v^2| + \sum_{uv \in E_G^{dif}} w_G(u, v) (x_u^2 + x_v^2)}{x^T L_H x} \\
&\geq \min \left\{ \frac{\sum_{u \in V^-, v \in V^-} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_u^2}{\sum_{u \in V^-, v \in V^-} w_H(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_H(u, v) x_u^2}, \right. \\
&\quad \left. \frac{\sum_{u \in V^+, v \in V^+} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_v^2}{\sum_{u \in V^+, v \in V^+} w_H(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_H(u, v) x_v^2} \right\}.
\end{aligned}$$

The inequality comes from applying of Lemma 1.

By symmetry in V^- and V^+ , it suffices to show that

$$\frac{\sum_{u \in V^-, v \in V^-} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_u^2}{\sum_{u \in V^-, v \in V^-} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_u^2} \geq \phi(G, H). \tag{3.2}$$

We sort the x_u in increasing order of $|x_u|$ into such that $x_{u_1} \geq \dots \geq x_{u_k}$, and let $S_k = \{x_{u_1}, \dots, x_{u_k}\}$. We have

$$\sum_{u \in V^-, v \in V^-} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_u^2 = \sum_{i=1 \dots k} (x_{u_i}^2 - x_{u_{i-1}}^2) \text{cap}_G(S_k, \bar{S}_k),$$

and

$$\sum_{u \in V^-, v \in V^-} w_H(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_H(u, v) x_u^2 = \sum_{i=1 \dots k} (x_{u_i}^2 - x_{u_{i-1}}^2) \text{cap}_H(S_k, \bar{S}_k).$$

Applying Lemma 1 we have

$$\frac{\sum_{u \in V^-, v \in V^-} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_u^2}{\sum_{u \in V^-, v \in V^-} w_G(u, v) |x_u^2 - x_v^2| + \sum_{u \in V^-, v \in V^+} w_G(u, v) x_u^2} \geq \min_k \frac{\text{cap}_H(S_k, \bar{S}_k)}{\text{cap}_H(S_i, \bar{S}_i)} \geq \phi(G, H),$$

where the second inequality is by definition of $\phi(G, H)$. This proves equation 3.2 and the Lemma follows. \square

We now proceed with the proof of the main Theorem.

Proof. We have

$$\begin{aligned} x^T L_G x &= \sum_{uv \in E_G} w_G(u, v) (x_u - x_v)^2 \\ &= \sum_{uv \in E_G^{\text{same}}} w_G(u, v) (x_u - x_v)^2 + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u - x_v)^2 \\ &\geq \sum_{uv \in E_G^{\text{same}}} w_G(u, v) (x_u - x_v)^2 + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2). \end{aligned} \tag{3.3}$$

The last inequality follows by $x_u x_v \leq 0$ as $x_u \leq 0$ for all $u \in V^-$ and $x_v \geq 0$ for all $v \in V^+$.

We multiply both sides of the inequality by

$$\sum_{uv \in E_G^{\text{same}}} w_G(u, v) (x_u + x_v)^2 + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2).$$

We have

$$\begin{aligned} &\left(\sum_{uv \in E_G^{\text{same}}} w_G(u, v) (x_u - x_v)^2 + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2) \right) \\ &\cdot \left(\sum_{uv \in E_G^{\text{same}}} w_G(u, v) (x_u + x_v)^2 + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2) \right) \\ &\geq \left(\sum_{uv \in E_G^{\text{same}}} |x_u - x_v| |x_u + x_v| + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2) \right)^2 \\ &= \left(\sum_{uv \in E_G^{\text{same}}} |x_u^2 - x_v^2| + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2) \right)^2. \end{aligned}$$

Furthermore, notice that $(x_u + x_v)^2 \leq 2x_u^2 + 2x_v^2$ since $2x_u^2 + 2x_v^2 - (x_u + x_v)^2 = (x_u - x_v)^2 \geq 0$. So, we have

$$\begin{aligned} &\sum_{uv \in E_G^{\text{same}}} w_G(u, v) (x_u + x_v)^2 + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2) \\ &\leq 2 \left(\sum_{uv \in E_G^{\text{same}}} w_G(u, v) (x_u^2 + x_v^2) + \sum_{uv \in E_G^{\text{dif}}} w_G(u, v) (x_u^2 + x_v^2) \right) \\ &= 2x^T D x \leq 4x^T L_{D_G} x, \end{aligned}$$

where D is the diagonal of L_G and the last inequality comes from Lemma 2. Combining the last two inequalities we get:

$$\frac{x^T L_G x}{x^T L_H x} \geq \frac{1}{2} \cdot \left(\frac{\sum_{uv \in E_G^{same}} |x_u^2 - x_v^2| + \sum_{uv \in E_G^{dif}} w_G(u, v)(x_u^2 + x_v^2)}{x^T L_H x} \right) \cdot \left(\frac{\sum_{uv \in E_G^{same}} |x_u^2 - x_v^2| + \sum_{uv \in E_G^{dif}} w_G(u, v)(x_u^2 + x_v^2)}{x^T L_D x} \right).$$

By Lemma 3, we have that the first factor is bounded by $\frac{1}{2}\phi(G, H)$ and the second factor bounded by $\frac{1}{2}\phi(G, D_G)$. Hence we get

$$\frac{x^T L_G x}{x^T L_H x} \geq \frac{1}{4} \phi(G, H) \phi(G, D_G). \quad (3.4)$$

□

4 Computation

Note that in Lemma 3 we actually proved that for any vector x such that $x^T d = 0$, we can sort x and find $n - 1$ sets $S_i \subseteq V$ such that:

$$\min_i \frac{cap_G(S_i, \bar{S}_i)}{cap_H(S_i, \bar{S}_i)} \leq \frac{1}{\phi(G, D_G)} \cdot \frac{x^T L_G x}{x^T L_H x}.$$

This suggests that we can find a cut which is at most $1/\phi(G, D_G)$ larger than the ratio $(x^T L_G x / x^T L_H x)$.

Given any positive definite matrix A , one can use the inverse power iteration $y_{i+1} = A^{-1}y_i$, where y_0 is a random vector, to find a vector x such that

$$\frac{x^T A x}{x^T x} \leq (1 + \epsilon) \lambda_{\min}(A). \quad (4.1)$$

The number of rounds required for this is $O(\log n/\epsilon)$; for a proof see [ST14].

Analogously, given a pair of positive definite matrices (A, B) , one can perform power iteration with the matrix $A^{-1}B$ to find a vector x such that

$$\frac{x^T A x}{x^T B x} \leq (1 + \epsilon) \lambda_{\min}(A, B).$$

The proof is similar to the simple eigenvalue problem case, using only the additional fact that the generalized eigenvectors of the pair (A^{-1}, B^{-1}) are the usual eigenvectors of the matrix $A^{-1}B$, in addition with the fact that the eigenvectors are A -orthogonal and B -orthogonal [SS90]. Note that the iteration can be extended to the case when A has a known null space, by simply operating on vectors orthogonal to the null space.

Note now that $A^{-1}B y_i$ can be implemented as a linear system solve $Az = B y_i$. Instead of solving exactly a linear system with the Laplacian, one can use a more efficient iterative solver instead, and ask for a solution \tilde{z} that satisfies $\|\tilde{z} - z\|_A \leq (1 + \epsilon/4) \|A^{-1}y_i\|_A$. For many iterative linear system solvers, including the fast Laplacian solvers, this step is an implicit multiplication with a matrix \tilde{A}^{-1} which is spectrally close to A^{-1} . Spielman and Teng [ST14] observe that this is sufficient for the computation of an approximate eigenvector that satisfies inequality 4.1. This extends to the generalized problem with Laplacians as well. Finally, a little more care has to be taken for the case of Laplacian solvers that are randomized. In that case $O(\log(1/p))$ different runs of the inverse power method are need to get a good approximate eigenvector with probability at least $1 - p$.

References

- [Chu97] F.R.K. Chung. *Spectral Graph Theory*, volume 92 of *Regional Conference Series in Mathematics*. American Mathematical Society, 1997.
- [Mih89] Milena Mihail. Conductance and convergence of markov chains-a combinatorial treatment of expanders. In *30th Annual Symposium on Foundations of Computer Science, Research Triangle Park, North Carolina, USA, 30 October - 1 November 1989*, pages 526–531. IEEE Computer Society, 1989.
- [SS90] G.W. Stewart and Ji-Guang Sun. *Matrix Perturbation Theory*. Academic Press, Boston, 1990.
- [ST14] Daniel A. Spielman and Shang-Hua Teng. Nearly-linear time algorithms for preconditioning and solving symmetric, diagonally dominant linear systems. *SIAM J. Matrix Anal. & Appl.*, 35(3):835–885, 2014.
- [Tre13] Luca Trevisan. Is cheeger-type approximation possible for nonuniform sparsest cut? *CoRR*, abs/1303.2730, 2013.